**Statement of Dr. John H. Marburger, III**
**Director, White House Office of Science and Technology Policy**
**to the Committee on Science, U.S. House of Representatives**
**on *H.R. 4218, the High Performance Computing Revitalization Act of 2004***
**May 13, 2004**

Mr. Chairman and members of the Committee, I am pleased to meet with you today to discuss high performance computing and share with you the Administration's views on the *High-Performance Computing Revitalization Act of 2004*. Networking and information technology (IT) research and development (R&D) continues to be one of this Administration's highest interagency R&D priorities, and the Office of Science and Technology Policy (OSTP) is actively engaged in interagency coordination of this area.

Advancements in IT underlie many of the most important technological developments of our time. The influence of IT is truly pervasive, having a profound impact on the way we work, learn, do business, and communicate. IT plays an enabling role in all of the President's priorities: winning the war on terrorism, securing the homeland, and strengthening the economy. Its impact in this last area has been particularly profound, with tremendous increases in productivity, in particular, serving to reshape the economy. Virtually all aspects of commerce today have felt the impact of IT, from product development to supply-chain management. Federally-funded R&D underpins these advances.

<u>The NITRD program</u>

For all of these reasons, the multi-agency Networking and IT R&D (NITRD) program, which represents the Federal government's combined R&D efforts in this field, has been and remains a priority of this Administration. As such, it has been featured in each of President Bush's budget requests to Congress. The R&D aspects of the Budget are in turn shaped in part by the memorandum that the Office of Management and Budget (OMB) Director and I send to the heads of agencies with science and technology responsibilities every year, outlining our top multi-agency R&D priorities. Agencies take this memo into account when crafting their budget submissions. The commitment to the NITRD portfolio signaled in these memos is reflected in the funding increases this program—one of the more mature R&D programs in the Federal portfolio—has realized. The increases to the NITRD portfolio total 14 percent, to over $2 billion, since President Bush took office in 2001.

A formal interagency working group, which exists under the National Science and Technology Council's (NSTC's) Committee on Technology, coordinates interagency efforts related to the NITRD program. The NSTC is a Cabinet-level council that advises the President on science and technology. It is chaired by the President or Vice President, though that responsibility is typically delegated to the OSTP Director. It is the principal means to coordinate science and technology matters within the Federal research and development enterprise.

The Interagency Working Group on NITRD is made up of experts from 12 different agencies with responsibilities for R&D in networking and IT. The group meets regularly and has established seven reporting categories in order to focus on particular areas of emphasis within the overall NITRD portfolio. These Program Component Areas (PCAs) cover the following areas: (1) high-end computing infrastructure and applications, (2) high-end computing research and development, (3) human computer interaction and information management, (4) large-scale networking, (5) software design and productivity, (6) high-confidence software and systems, and (7) social, economic and workforce issues related to IT. Coordinating groups associated with these PCAs meet regularly to determine research needs, coordinate activities, and review progress.

Every year, the NITRD "blue book"—a supplement to the President's Budget—outlines the activities and funding levels for each of the seven areas listed above. This document provides more detailed descriptions of NITRD program activities and more specific budgetary information than is present in the overall Budget. The FY 2005 blue book will be available this summer.

The President's Information Technology Advisory Committee (PITAC), which is made up of private sector representatives with expertise in IT, provides expert, outside advice to the NITRD program. President Bush announced his intention to appoint the current 24 members of PITAC to their positions in May of last year. They have since tackled the important issue of the role of IT in the health care system, and are embarking on an examination of the Nation's cybersecurity R&D activities. A future activity will address issues related to computational science, a field that focuses on scientific simulation.

## The *High-Performance Computing Revitalization Act of 2004*

Both the NITRD program's and PITAC's foundations are found in the *High Performance Computing Act of 1991* (P.L. 102-194). The Act, which was subsequently updated with the *Next Generation Internet Act of 1998* (P.L. 105-305), defines an interagency program for the Nation's networking and IT R&D activities. It required the formation of goals and priorities for high-performance computing, which was defined broadly to mean "advanced computing, communications, and information technologies…" It required establishment of an advisory committee to provide outside advice to the program, and identified specific agency activities.

The program that developed from this legislation—the NITRD program—is flourishing today. In the *High-Performance Computing Revitalization Act of 2004,* the Committee has provided a timely update of this important legislation while preserving the original legislation's intent and scope. I share your enthusiasm for and commitment to high-performance computing and I am pleased to convey the Administration's support for the *High-Performance Computing Revitalization Act of 2004,* in its current form.

## High-end computing within the NITRD program

High-end computing—or supercomputing, as it is sometimes referred to—is an important element of the NITRD program. Certain of today's important and unsolved scientific and engineering problems can be answered only with high-end computers employing hundreds to thousands of times more computational power than is available in today's systems. These unsolved problems include important national security challenges in areas such as cryptanalysis and image processing of satellite and other data, as well as important scientific and technological questions related to the analysis of complex systems such as aircraft, the atmosphere, and biological systems.

Two PCAs exist to support interagency coordination of high-end computing within the NITRD program, one on Infrastructure and Applications, and the other on R&D. Together, they encompass advances in hardware, software, architecture, and application systems; advanced concepts in quantum, biological, and optical computing; algorithms for modeling and simulation of complex physical, chemical, and biological systems and processes; and information-intensive science and engineering applications.

A number of agencies with active interest in high-end computing participate in coordination: the National Science Foundation (NSF), the National Institutes of Health (NIH), the National Aeronautics and Space Administration (NASA), the Department of Defense (DOD), which includes the Defense Advanced Research Projects Agency (DARPA), the National Security Agency, and the Office of the Director, Defense Research and Engineering, the Department of Energy (DOE) (both the Office of Science and the National Nuclear Security Administration), the National Institute of Standards and Technology (NIST), the National Oceanic and Atmospheric Administration, and the Environmental Protection Agency (EPA).

High-end computing has been and continues to be a high-priority area within the NITRD program. The President's FY 2004 and 2005 Budgets stressed the importance of high-end computing, as did the OSTP/OMB FY 2005 guidance memorandum I referred to earlier.

NSF's and DOE's provision of high-end computing resources to academic researchers

I understand that the Committee is particularly interested in better understanding the provision of high-end computing resources by DOE and NSF to university researchers. NSF remains the largest provider of supercomputing resources to academic researchers, though need continues to outstrip demand. In addition to NSF-funded scientists and engineers, users include large numbers of NIH-, NASA-, and DOE-funded scientists and engineers.

NSF support for high-performance computing will continue to advance a broad range of science and engineering areas, with emphasis on the support of university-based science and engineering research and education. Moreover, the national community has identified a pressing need to create a state-of-the-art cyberinfrastructure that integrates and makes broadly accessible state-of-the-art high-performance compute nodes, research instruments that generate research data, data storage and management resources, visualization tools that advance capabilities to interpret and analyze data, and new tools for collaboration.

Responsive to this need, NSF's focus on cyberinfrastructure will continue to advance high-performance computing while broadening the scope of facilities and services supported to create new science and engineering knowledge. In addition, NSF will continue, through education, outreach and training as well as development of "services" to make this new cyberinfrastructure available to and usable by a wider range of the national research and education community.

NSF-funded high-performance computing centers include the San Diego Supercomputing Center, the National Center for Supercomputing Applications, and the Pittsburgh Supercomputing Center. These Centers are partnering in the Teragrid effort that integrates their leading edge high-end computing facilities with complementary resources at the California Institute of Technology, Argonne National Laboratory, Indiana University, Purdue University, the University of Texas, and Oak Ridge National Laboratory; the resources are connected by a high-performance backbone network (40 gigabytes/second). NSF's Middleware Initiative is developing software to support distributed applications including collaboration and grid computing.

NSF builds on a wide range of collaborations among universities, federal partnerships (including DOE and DOE Labs), and other sectors. Access to these facilities is available to university researchers through application to the centers. Accounts tailored to development, mid- and high-range needs, educational use, and for Southeastern Universities Research Association and Experimental Program to Stimulate Competitive Research applicants are available. The Partnerships for Advanced Computational Infrastructure and Teragrid facilities allocated more than 169,000,000 CPU (central processing unit) hours to users in FY 2003. Upgrades, both in progress and planned, will significantly increase available CPU hours.

NSF continues significant investments in high-end computing; NSF plans $70 million in FY 2005 for high-end computing facilities. This investment is complemented by significant investments in education, outreach and training, which increase the number and diversity of the user communities, as well as investments in application codes, software, and new technologies for the next generation of computing.

DOE's Office of Science operates several high-end computing facilities, including (1) the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory, which is the flagship high-end computing facility for the Office of Science; (2) the Center for Computational Sciences (CCS) at the Oak Ridge National Laboratory; and (3) the Environmental Molecular Sciences Laboratory (EMSL) at Pacific Northwest National Laboratory. All are managed as unclassified open facilities in support of the DOE Office of Science mission. University researchers who are working on applications that are relevant to the broad science mission of the Office of Science can apply for access to these facilities, which is granted on a competitive peer-reviewed basis. For example, up to seven percent of NERSC resources are available to researchers for mission-relevant work that is not directly supported by the Office of Science.

An exception to the requirement for mission relevance is DOE's *Innovative and Novel Computational Impact on Theory and Experiment* (INCITE) program at NERSC. The goal of the program is to provide ten percent of the computational resources at NERSC in very large allocations to a small number of computationally intensive large-scale research projects selected based on their ability to make high-impact scientific advances. The INCITE program specifically encouraged proposals from universities and other research institutions.

In FY 2004, 52 proposals were submitted, with more than 60 percent coming from academic researchers, requesting a total of more than 130 million hours of supercomputer processor time. The three awards in FY 2004 amount to ten percent of the total computing time available on NERSC's current IBM supercomputer.

The Office of Science yesterday announced an award for their "Leadership-class System," a $25 million investment in FY 2004. The request for applications for acquisition of this leadership-class system specified that "Proposed activities should be designed to support computational science applications research areas relevant to the mission of the Office of Science, as well as those of other federal agencies." University researchers—regardless of which Federal agency supports their work—will be granted access to this leadership-class computational resource, again on a competitive peer-reviewed basis.

<u>Challenges facing the high-end computing enterprise</u>

The challenges facing high-end computing today are significant. Decisions made years ago—sensible at the time—led to a dependence largely on bundled clusters of commercial-off-the-shelf (COTS) processors. The promise of high aggregate performance at relatively low cost made the choice of these systems highly attractive. However, we now know that while these systems are effective for some classes of applications, many others—including certain applications relevant to national security considerations—are poorly served by COTS-based solutions. Addressing this problem, however, is costly—prohibitively so—for all but a few Federal agencies and virtually all private-sector enterprises.

In the 1990s, due to the limited market for high-end computing systems and the dramatic expansion of the market for low and mid-range systems, the U.S. computer industry focused primarily on the hardware and software needs of business applications and smaller scale scientific and engineering problems. As a result, the flow of R&D needed to maintain high-end computing technologies in the U.S., and the human capital required to sustain its cutting edge, have failed to keep up with opportunities for development.

<u>The High-End Computing Revitalization Task Force</u>

With these concerns in mind, OSTP initiated the organization of a task force, under the auspices of the NSTC, made up of agency experts in high-end computing. This High-End Computing Revitalization Task Force (HECRTF) was given a specific charge based on the issues outlined in the President's FY 2004 Budget, which said:

> *"Due to its impact on a wide range of federal agency missions ranging from national security and defense to basic science, high-end computing—or supercomputing—capability is becoming increasingly critical. Through the course of 2003, agencies involved in developing or using high-end computing will be engaged in planning activities to guide future investments in this area, coordinated through the NSTC. The activities will include the development of an interagency R&D roadmap for high-end computing core technologies, a federal high-end computing capacity and accessibility improvement plan and a discussion of issues (along with recommendations where*

*applicable) relating to federal procurement of high-end computing systems. The knowledge gained from this process will be used to guide future investments in this area. Research and software to support high-end computing will provide a foundation for future federal R&D by improving the effectiveness of core technologies on which next-generation high-end computing systems will rely."*

Specifically, the Task Force was asked to develop a forward-looking plan for high-end computing with the following three components: (1) an interagency R&D roadmap for high-end computing core technologies, (2) a federal high-end computing capacity and accessibility improvement plan, and (3) recommendations relating to federal procurement of high-end computing systems.

I am pleased to provide the Committee with the Task Force's report, the *Federal Plan for High-End Computing*. In its report, the Task Force addresses the needs of major Federal science and technology areas for high-end computing, articulating and synthesizing the urgent problems facing high-end computing.

The Task Force lays out detailed roadmaps for investments in key R&D areas, which include hardware, software, and systems. They emphasize the importance of addressing the increasing gap between the theoretical peak performance and the sustained system performance of high-end computers—a problem that has plagued the massive multiprocessor systems currently in use. Their report also emphasizes the need for procurement of "early access" systems that will enable the development of more robust systems and help identify failed approaches before full-scale procurements take place.

The report also addresses issues related to the acquisition, operations, and maintenance of high-end computing systems by agencies, including so-called "leadership class" systems—leading-edge, high-capability computers capable of tackling heretofore unsolvable computational problems. The Task Force recognized that the costs associated with the development of leadership systems are beyond the reach of almost any agency working alone. At the same time, the Task Force emphasized that the need is great: demand for high-end computing capabilities surpasses the resources available in every agency, and some of the smaller agencies, such as EPA and NIST, rely on the resources of other agencies to meet their need. To address this, the Task Force recommends that future leadership systems be treated as national resources, for use by all of the agencies that participate in the system's development (and those agencies' constituents). They suggest specific mechanisms by which agencies that lack the resources to develop high-end computing systems can partner with larger agencies for access to existing systems.

Additional sections of the report address procurement issues, which are currently hampered by the diversity of agency needs for high-end systems and their practices governing procurement of them. The Task Force suggests the initiation of several pilot projects related to procurement to address this. These include the development of improved suites of benchmarks that better mirror applications, an evaluation of the total cost of ownership of several similar systems, and the development of a common solicitation and use of a single suite of benchmarks for procurement, using lessons learned from the first two pilot projects.

Finally, the report describes interagency mechanisms through which to coordinate implementation of various aspects of the plan.

It is important to recognize that benefits of the Task Force's work have already begun to accrue, with the high level of interagency cooperation already leading to tangible results. For example, three agencies—NSF, DOE's Office of Science and DOD—have combined forces to initiate the High-End Computing University Research Activity, a pilot program aimed at funding basic research in different "theme" areas related to high-end computing. Joint planning has led to two closely coordinated solicitations. With software as the theme for 2004, NSF recently issued a program solicitation (that also incorporates DARPA interests) for research on "Software and Tools for High-End Computing." This program, for which the anticipated funding of $7 million was provided by both NSF and DARPA, will support "innovative research activities aimed at building complex software and tools (on top of the operating system) for high-end architectures." A second solicitation, from DOE's Office of Science but also with DARPA interest and funding, is focused on "Operating/Runtime Systems for Extreme Scale Scientific Computation." The agencies' involvement in the HECRTF was a key factor in the development of these programs, and a sign of the future benefits we can expect from this important effort.

I commend the Task Force for developing their report and for their commitment to continue the work that they have begun by making high-end computing a continued, vigorous interagency activity that fully captures the synergies evident in their report. I look forward to working with all of the agencies this year to see that the Task Force's recommendations are considered in the preparation of agencies' FY 2006 budget requests. Addressing the issues facing the Nation's high-end computing enterprise will require a sustained and coordinated effort. The Task Force's report constitutes an important first step.